The Economics of Transportation Networks

By Yuri V. Yevdokimov
Department of Economics
University of New Brunswick
P.O. Box 4400, Fredericton, NB E3B 5A3
Phone: (506) 447-3221
E-mail: yuri@unb.ca

Abstract

The majority of studies in the field of transportation economics usually ignore the fact that transportation is a group of multiple services to multiple users produced and consumed in a transportation network. Although some of these studies do recognize the importance of transportation networks, they narrowly define them, missing some of their rather important features.

In this paper, transportation network is analysed from a standpoint of the literature on economics of network. According to this literature, transportation network is defined as a system that consists of transportation infrastructure, transportation vehicles and some industry specific services to perform a function of moving people and freight. Then production and consumption of transportation services are analysed in the light of the network effects, discussed and developed in the literature on economics of networks.

As well, the systems approach is applied to model transportation. Basic principles of the approach with respect to transportation are formulated. According to these principles, transportation network is incorporated in an economic system that consists of three levels, micro, meso and macro. All three levels are then connected in order to optimize the whole system and find optimal characteristics of transportation as a component of the system. Finally, computer simulation exercise, based on the designed model, is performed and its results are discussed.

Introduction

It appears to be that for many years transportation economics has been dominated by two approaches. The first one, being applied microeconomics, regards transportation either as an input in general production process in the case of freight transportation or as a service in a household consumption basket in the case of passenger transportation. Consequently demand for transportation is derived either as a factor demand or as a Marshallian demand. There exists a variety of demand models based on this framework.

Since transportation has been regulated until 1980s, supply side analysis of transportation under the first approach has not been as popular as demand side. After deregulation, however, the supply side of transportation was given much more attention. In general, supply side analysis is based on specification of the cost structure of transportation service provider. Usually it is assumed that transportation cost curves resemble theoretical shapes discussed in standard microeconomic textbooks.

Overall in this approach, transportation is a subject to a standard microeconomic analysis when a specific market is defined as to make transportation service under study homogeneous to be able to apply concepts of the neo-classical microeconomic theory.

In the second approach, transportation is viewed through transportation networks. Although transportation networks have existed ever since a road connected two cities, economic studies of these networks are a relatively recent phenomenon. There are different meanings that economists have historically incorporated into the concept of a transportation network. Roughly it is possible to group them into three types: (i) network as infrastructure; (ii) network as transportation cost function; and (iii) network as a spatial object.

In this approach, each of the three interpretations of a transportation network emphasizes just one out of several important features of the transportation network, on the one hand, and spatial nature of the networks is interpreted in a geographical or topological sense, on the other.

Hence, let us first analyse general deficiencies that arise as a result of the two conventional approaches to the economics of transportation, deficiencies that call for a different framework.

Conventional economics of transportation

The majority of transportation economists still treat transportation as a market. However, let us take a closer look at our traditional, microeconomic interpretation of a market. There are three basic conditions that should be satisfied:

1. Homogeneous good or service
2. Many buyers (service users) and one, few or many sellers (service providers)
3. Private ownership, private good or service

Applied to transportation, we find that this framework violates at least two of the above three conditions. First of all, transportation is a group of heterogeneous services with different attributes. Second, transportation services are produced and consumed in transportation networks with some network pieces publicly owned. Therefore, in order to treat transportation as a market, it is necessary to narrow this market in such a way as to artificially make the service under study a quasi-homogeneous. For example, the market for *air passenger transportation in Atlantic Canada*. But still, we would violate condition number three since airports or the so-called transportation fixed facilities are usually publicly owned.

And how about the other approach when transportation is viewed in the light of transportation networks? There is no assumption of homogeneity of transportation services in this approach.

However, as already noted, each of the three interpretations of a transportation network emphasizes just one out of several aspects of the transportation network. Let us discuss all three in detail.

Studies examining transportation networks as infrastructure were poplar in the United States in the 1950s and 1960s (Cain, 1997). This view on transportation networks is still popular in economic literature on the role of public infrastructure in economic growth and development (see Ferrara and Marcellino, 2000 for review of this literature). Under this view, infrastructure was narrowly defined as public capital stock that includes only tangible, non-military capital goods. Later it was re-defined as "large physical capital facilities and organizational, knowledge, and technological frameworks that are fundamental to the organization of communities and their economic development" (Tatom, 1993). Even though the latter definition is broader than the former, nevertheless in the way it is defined, it still fails to capture some important features of a transportation network, discussed later in this paper, and first of all its spatial nature. As well, the definition reflects only a fraction of the total value of the transportation network.

Since the 1960s the view on transportation networks has become wider. Many macroeconomic models of regional development as well as general equilibrium models were designed in which an economic system was presented as a set of sectors or markets that interacted with each other. Transportation was introduced into these models as well (see, for example, Lipsey and Steiner, 1969 and Roson, 1995). Transportation network in these models is represented by the cost of transportation – a linear function of distance between demand and supply areas for various commodities. Like the previous approach, this one fails to capture the spatial nature of the network. As well it cannot explain what happens to the transportation cost if distance between demand and supply areas is unchanged but a new link is added (say, one more air route between already existing service points) or an old link is improved (say, a four-lane highway instead of a two-lane highway).

The third interpretation of a transportation network is associated with the so-called spatial input-output models that regard the network as a spatial object (see, for example, Rohr and Williams, 1994). However, the spatial nature of the transportation network in these models has cartographical, topological interpretation when co-ordinates and geographical location matter which requires massive databases. Although this type of models does reflect the spatial nature of transportation networks, nonetheless it interprets it in a way different from modern literature on networks.

The existing literature on networks emphasizes that spatial should denote the dimension or size of the network components, not geographical locations. That is why in this literature, a network is viewed as (i) existing capacity, and (ii) capacity utilization. For example, the following fundamental network measures are found in this literature:

Gamma Index: Compares the actual number of links with the maximum possible number of links

Alpha Index: Compares the number of actual "circuits" (traffic) with the maximum number of all possible "circuits"

Beta Index: Compares the number of links with the number of nodes (service points)

Therefore, it appears to be that both conventional approaches do not capture the real nature of transportation as a group of multiple services to multiple users, produced and consumed in transportation networks – networks in terms of capacity and capacity utilization. As well, the two conventional approaches can not be used if one wants to model and explain such transportation phenomena as sustainable transportation, intermodal transportation, intelligent transportation systems and other.

Hence, we need a different framework, and it appears to be that such a framework can be based on the emerging literature on economics of networks which is discussed next.

Transportation as a network industry

Nicholas Economides in his seminal paper *The Economics of Networks* (1996) wrote that *the modern economy would be very much diminished without the transportation, communication, information, and railroad networks.* As we can see from the above citation, one of the most prominent specialists in the field of economics of networks includes transportation as a subject matter of his analysis into the networks.

The field of the economics of networks has been rapidly developing during 1990s. Although mostly it is associated with the development of information technology, all of its fundamentals can be applied to transportation network as well. Actually the main goal of this paper is application of these fundamentals to study consumption and production of transportation services in transportation networks.

Let us begin this analysis by presenting a definition of a network, introduced by Katz and Shapiro (1994) – the other two leading scholars in the field. They defined network as *a system of compatible devices, a system that can be any combination of a durable good and associated goods and services that perform some desired function*. In terms of transportation, the desired function is to move people and freight, a durable good is transportation fixed facilities, associated goods are transportation vehicles and associated services are labour services of vehicle operators, maintenance of fixed facilities, traffic operation, policing and some others. In an economic sense, the above definition represents definition of the production function for transportation services as a function of the following inputs: (i) transportation immobile capital: service points with links that connect them (e.g. railways, terminals, highways, airports, seaports, pipelines); (ii) transportation mobile capital:

transportation vehicles (e.g. cars, trucks, aeroplanes, trains, ships); (iii) services (e.g. labour services of transportation vehicle operators).

Here it is necessary to emphasize the term *system* used by Shapiro and Katz in their definition of a network. As a matter of fact, system means that it is represented by elements whose collective behaviour is essential to the whole, but not relevant by themselves. However, at this point it is necessary to separate two specifications of a transportation network as a system: (i) transportation network as a physical system, and (ii) transportation network as an economic system.

Transportation network as a physical system is the subject matter of an engineering analysis. In the engineering sense, transportation system is characterized by infrastructure (say roads) and vehicles (say cars) as the system's components that produce and carry out flows (traffic). In this paper however, the proposed framework regards transportation as a part of an economic system, a system consisting of the transportation network as defined earlier, the network users, transportation service providers and the rest of the economy. The chart presented in Fig. 1 illustrates the framework.

The chart shows three levels: micro, meso and macro. At level zero, basic assumptions about economic system are made (Block 0). These assumptions affect the functioning of the entire economic system. Transportation network is defined at the micro level. At this level, analysis focuses on behaviour of the network users and providers of transportation services (Blocks 1 and 2) with respect to the pre-specified assumptions (Block 0). As a result, some optimal choices are made which affects other sectors (markets) of the economy. Hence, at meso level the changes in the affected sectors are analysed (Block 3). Broken lines connecting Blocks 1,2 and 3 reflect the iterative nature of the process. The process stops when all relevant markets are in equilibrium. At macro level,

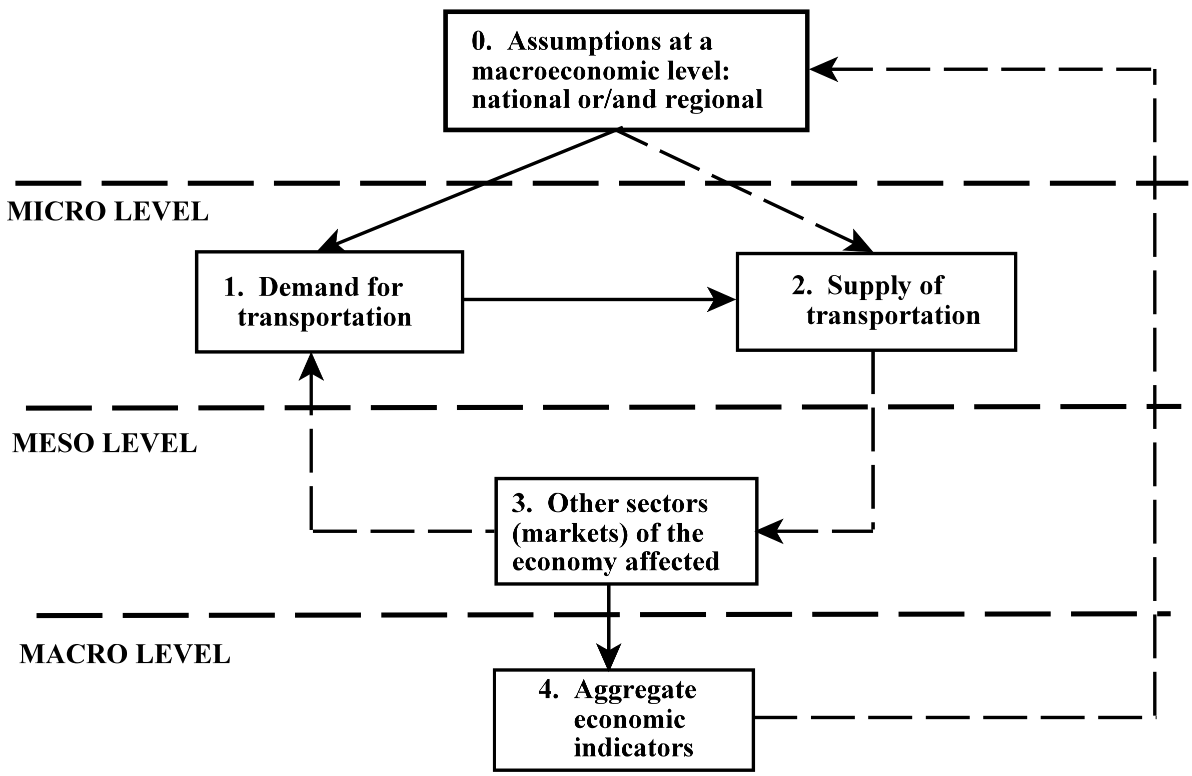

Fig 1. Systems Approach to Modelling Transportation Network

economic indicators from micro and meso levels are aggregated (Block 4). Then the aggregate indicators are compared with the initial assumptions (Block 0). If the assumptions and the obtained macroeconomic indicators contradict each other, the assumptions are changed, and the process is repeated, which is reflected by a broken line connecting Blocks 4 and 0. In a sense, the framework is very similar to the General Equilibrium Model (GEM). However, interpretation of the transportation sector at micro level is different which affects the entire system presented in Figure 1.

Let us define the *physical network* as the network's fixed facilities (service points with links or immobile transportation capital or infrastructure). According to the latter definition, production of transportation services is realized by service providers with the help of *transportation vehicles* and *labour services* in a *physical network* – all three being means of production. Consumption of transportation services occurs in the physical network as well, and as such, it is associated with the total value of the network. Moreover, in order to keep consistency with the existing classification, transportation services include both passenger and freight transportation.

Furthermore, as any network, the transportation network is subject to the network effects. In the literature these effects are known as *network economies* or *network externalities*. For instance, a pure transportation economist Boyer (1998) argues that network economies refer to the reduction in unit cost in a network as economic activity expands which is obviously a supply side (production) effect. On the other hand, the most of the literature in the economics of networks (see, for example, Page and Lopatka, 1997) defines network economies as demand side (consumption) externalities. Even though many scholars in the field of the economics of networks do not distinguish between the network economies and the network externalities, we will follow Katz and Shapiro (1994), who argue that network economies (or network effects) are more than just demand side externalities.

Network effects in a transportation network

In the literature, a network externality is defined as a benefit conferred on users of a product by another's purchase of the product (Page and Lopatka, 1997). In other words, network externalities exist when the value of a product or service to a user is affected by the number of other users in the network. In this definition, it is explicit that network externalities are regarded as positive consumption externalities or demand-side externalities. However, in many industries, and transportation is one of them, the above defined network externalities are not the only source of the network effects, and they are not necessarily the most important. While network externalities are an important element of transportation as a network industry, economies of scale of specific nature are as important if not even more. Let us discuss the transportation network effects that arise on demand and supply side separately.

Demand side effects

It appears that for transportation to be of any use it must occur in a network since consumers of transportation are physically connected to the transportation network. Therefore, the value of transportation services to the consumers is directly associated with the value of the transportation network as a whole. This is how consumption network externalities come into play.

As noted, network externalities imply that the value of a product or service increases as the number of users of the product or service grows (Page and Lopatka, 1997). In other words, the value of the product or service produced in a network depends on the size of the network. With respect to transportation, when the size of a physical network increases, a user of the network receives extra benefits in the form of a wider accessibility to new locations. These are direct positive externalities that arise as a result of the physical network expansion which increases the number of the network's users and consequently the network's value.

There is a number of indirect transportation network externalities as well. Indirect externalities arise if increased economic activity, surrounding the existing physical network, results in more options and/or better service for current users. We know that a well developed transportation network attracts businesses to relocate into the area which eventually increases the value of transportation services produced in the network. For instance, increased number of motels and restaurants along a highway makes one's trip more pleasant. Sometimes a well developed transportation network leads to the so-called economies of agglomeration – positive effects that arise due to concentration of economic activity. In some cases, the concentration of economic activity around well developed transportation network may even lead to the physical agglomeration of a number of firms engaged in a similar activity known as industrial clusters with positive externalities and knowledge spillovers (Krugman, 1995).

Examples of the transportation network externalities are enormous: population growth, introduction of a new transportation service (e.g., new bus route), expansion of the network through addition of a new service point and/or link (e.g., new highway, new air route), improvement in quality of existing transportation service (e.g., increase in frequency of the existing service, just-in-time delivery).

In all cases we end up with an increase in the number of the network users which increases the total value of the network according to the so-called Metcalfe's Law: *the value of a network of size* $\boldsymbol{n}$ *is proportional to* $\boldsymbol{n}^2$. In this definition, $\boldsymbol{n}$ stands for the number of the network users. Obviously enough that introduction of a new transportation service into the network attracts new users, on the one hand, and widens the choice available to current users, on the other. Addition of a new service point and/or link, while directly increasing the number of users, results in accessibility to new locations by current users. Improvements in quality of transportation services also attract new users while making the services more desirable for current users.

Hence, like in any network industry, demand for transportation services depends not only on price of transportation, but also on the number of current users. Number of current users is defined in terms of the number of consumers at a point in time. Consequently, whatever increases the number of current users will shift the demand for transportation, produced in the network under study, to the right reflecting positive network externalities.

There is one more shift parameter in the demand function - expectations about the network's size. Expectations about increase in future economic activity would increase the number of the network's users today. For instance, if businesses expect higher economic activity in a specific area, they will immediately relocate there creating positive externalities as follows: (i) the value of the transportation network increases directly since new businesses are new consumers of transportation; and (ii) the value of the network increases to all current users in the form of higher variety of products and services available to them.

On the other hand, a well developed transportation network implies high quality accessibility to a variety of locations. In an engineering language, it has sufficient capacity to accommodate current and potential users over time. As well, it assumes a relatively high capacity utilization of the network as a result of the ongoing increase in the number of the network's users. Hence, it is possible to identify three basic determinants of the network's size as defined by the number of current users – the network's capacity, capacity utilization and population at service points. For example, introduction of a new service or improvement in quality of existing service increase capacity utilization of the existing transportation network while addition of a new service point and/or link increases the network's capacity. Both lead to increase in the number of the network's users and consequently to a shift in the demand for transportation. Capacity of the transportation network can be defined in terms of maximum possible flow of transportation services that can be generated by transportation fixed facilities. Capacity utilization is the ratio of current flow of transportation services to the maximum possible flow which is the *alpha index* defined earlier. Such a framework gives us a nice feedback from production side which is what we intended to do – connect both sides of the transportation network to better model transportation. In addition, population growth directly shifts the demand for transportation over time.

So, demand for transportation is the relationship between volume of transportation services demanded and three variables: (i) the price of transportation (eg., freight rate, airfare); (ii) the number of the network users, and (iii) the expected number of the network users. Moreover, the latter two depend on the network's capacity and capacity utilization.

Supply side effects

Traditionally, networks were analysed under the assumption that each network was owned by a single firm (Economides, 1996). Thus, economic research focussed on the efficient use of the network structure as well as on the appropriate allocation of costs (see Sharkey, 1993). However, technological changes in the 1980s and 1990s eliminated the case of a monopoly in some of network industries as the only possible and other models were introduced[1] .

With regard to transportation, transportation services are produced by different service providers in a physical network. In almost all cases transportation vehicles are owned by a private service provider. However, transportation fixed facilities are sometimes owned by a private service provider, while in other cases the facilities are publicly owned. Hence, it appears to be that in some circumstances a private service provider owns the entire network and a transportation service produced is known as the fully integrated service (Boyer, 1998). In other cases, several service providers share transportation fixed facilities which are owned by the government.

---

[1] For example, communication networks are now dominated by several long-distance service providers as opposite to the case of a single provider in the 1980s

Mixed ownership in a transportation network is not the only distinguishing feature of transportation. Physical capacity constraints and limited flexibility of the transportation capital, both immobile and mobile, are the other two. Compared to information networks with almost unlimited capacity, transportation fixed facilities and vehicles do have capacity constraints. These physical capacity constraints eventually lead to congestion that translates into an increasing average total cost of transportation.

Limited flexibility of transportation capital is known as lumpiness. Transportation capital is added in discrete amounts due to technological requirements rather than economic cost minimization. It results in initial excess capacity of transportation fixed facilities and vehicles, and eventually in non-convexities of total production cost.

Furthermore, network components are complementary to each other which is one of the basic features of a network (Economides, 1996). It implies that transportation fixed facilities and vehicles, or what is defined in transportation economics literature as transportation capital (see Button, 1993 and Boyer, 1998), are complements.

It appears to be that physical capacity constraints of transportation network and its components, limited flexibility of transportation capital, complementarity of the transportation capital lead to some specific scale effects. The concept of economies of scale in multi-service transportation network can be thought of as responsiveness of the network components to increase in traffic levels. Sometimes such an increase will require adjustment in all network components, however, sometimes only some, but not all network components need to be adjusted accordingly. For example, a ten percent increase in traffic on a highway does not require a proportional increase in the size of the highway if its physical capacity is not reached. On the other hand, such an increase can be a result of either increase in number of vehicles on the highway or frequency of vehicle utilization.

Actually, scale effects in the transportation network can be achieved through two channels – increase in the network's capacity utilization and increase in the network's capacity. Hence, there are two scale effects – the so-called economies of density and economies of size. Economies of density refer to the situation when average total cost decreases with increase in traffic level due to increase in capacity utilization of transportation capital, vehicles and fixed facilities. In turn, economies of size are associated with decreasing average total cost with increase in traffic due to increase in physical capacity.

For example, a scheduled cargo consolidation and containerization at intermodal facility allows a service provider, a railway company to increase length of a train and frequency of service. It results in an increase in capacity utilization of both, vehicles (trains) and fixed facilities (railways). Passenger consolidation at airports allows a service provider, airline to increase load factor of existing planes. Increased frequency of intercity bus transportation leads to increase in capacity utilization of a highway. All of the above are examples of the increase in capacity utilization of the transportation network which may lead to economies of density.

On the other hand, addition of a road to the existing network, one more lane to the existing highway or addition of a new air route between two airports may lead to economies of size. It is so because all of the above is associated with expansion of the physical transportation network which increases its physical capacity.

Therefore, we define economies of density as a decrease in average total cost of transportation due to increase in capacity utilization of the existing transportation network. As well, it is possible to define diseconomies of density as an increase in average total cost due to increase in capacity utilization. The latter usually occurs if traffic levels approach the capacity constraint which is known as congestion. Consequently, we define economies of size as a decrease in average total cost of transportation due to increase in the network's capacity.

Previously we defined the production function for transportation services as the relationship between the volume of transportation and the following three factors of production: (i) transportation immobile capital, (ii) transportation mobile capital and (iii) services. Let us call it aggregate production function of the transportation network. According to our discussion, the function has to reflect the following basic features of the production of transportation services in the transportation network: (i) discrete nature of transportation capital; (ii) complementarity of transportation capital; (iii) changing capacity utilization; (iv) capacity constraints.

Supply of transportation can be obtained from the aggregate production function of the transportation network by deriving factor demand functions and substituting them into original production function. In such a setting, volume of transportation, supplied by a service provider, will depend on (i) the price of transportation, (ii) the network's capacity and (iii) the network's capacity utilization.

Illustration of the methodology

In order to apply the theoretical framework discussed in the paper, let us analyse a simple regional transportation network presented in figure 2:

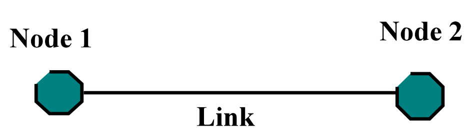

Fig. 2

This is a two-way transportation network consisting of two nodes (service points) and a link between the nodes. It is possible to think of the network as a highway connecting two cities. In fact, the network is a part of the regional economy. Let us apply the systems approach represented in Fig. 1.

0. Assumptions at macroeconomic level:

It is assumed that the regional economic system produces aggregate manufactured commodity Y and transportation is an input in the production of the aggregate manufactured commodity.

1. Demand for transportation:

There are two groups of consumers of transportation in the network – households and the manufacturing sector. Demand for transportation by households can be broken in two components: (i) T_H - quantity of commercial transportation (both freight and passenger) demanded by households; for simplicity it is assumed to be given exogenously; (ii) T_P - quantity of personal transportation or transportation by personal means of transportation; it is assumed to be given by the following function

$$T_P = a + bK_H + cY$$

where K_H is the number of personal cars; Y is the total output of aggregate manufactured commodity; a, b and c are parameters. The term cY captures positive indirect network externalities discussed earlier. In this case, indirect externalities can be thought of as an increase in the value of transportation, associated with expansion of business activity around the existing *physical* network.

Quantity of freight transportation demanded by the manufacturing sector T_M is given as a share of total output Y as follows

$$T_M = \alpha Y$$

where α is the transportation's share in total output. Hence, total volume of transportation in the network is[2]

$$T = T_H + T_P + T_M$$

2. Supply of Transportation:

For simplicity, a single provider of transportation services both freight and passenger is assumed. The provider owns commercial vehicles, but transportation infrastructure (immobile capital) is provided by the government. Supply of commercial transportation services in the network is based on the following cost minimization problem:

$$\min(w_1 K_C + w_2 L_C + w_3 K_F)$$

subject to

$$\frac{dK_F K_C^{\beta} L_C^{\gamma}}{u(T, K_F)} \geq T_H + T_M$$

where K_C is the number of commercial vehicles; L_C is the labour services of operators of commercial vehicles; K_F is the quantity of transportation immobile capital (fixed facilities or infrastructure); w_1, w_2 and w_3 are factor prices; d, β and γ are parameters. More about the production function of a transportation network can be found in Yevdokimov (2001). The function directly incorporates congestion in the network through capacity utilization u, the latter being a function of the transportation immobile capital K_F and traffic T. Since personal transportation T_P is the result of the personal production and consumption, its supply is the same as demand.

---

[2] The volume of transportation T can be thought of as traffic in the network, measured in terms of vehicle-kilometres per period

3. Other sectors (markets) of the economy affected:

Here it is assumed that the automobile sector, or the sector which produces vehicles, and the labour market are affected. However again, in order to keep things simple, this level of the economic system is presented by the following two relationships:

- automobile sector: $K_V = K_H + K_C$ where K_V is the total number of vehicles in the system;
- labour market: $L_C = L_C$.

4. Aggregate economic indicators:

The macro level is represented by the production function for aggregate manufactured commodity:

$$Y = AK^{\mu} L^{\nu} T^{\alpha}$$

where K is the manufacturing capital; L is the manufacturing labour; T is transportation; μ, ν and α are parameters.

According to the described economic system, it is possible to evaluate the following four scenarios:

Scenario 1. Private provider of transportation recognizes neither congestion nor positive network externalities. In terms of the chosen notation, it implies $a = 0$, $b = 0$, $c = 0$, $w_1 = 0$ and $K_F = fixed$.

Scenario 2. Private provider of transportation recognizes the existing congestion, but does not recognize positive network externalities or $c = 0$ and $K_F = fixed$.

Scenario 3. Both congestion and positive network externalities are taken into account under fixed K_F.

Scenario 4. Social optimum: congestion and positive network externalities are taken into account, and the physical network is allowed to adjust or K_F is variable.

In order to evaluate these scenarios, specific parameter values were chosen, and a computer simulation exercise was performed using MATHCAD8 program. Below the results of the computer simulation exercise are presented.

Table 1. Results of computer simulation

Indicator	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Commercial vehicles, K_C	10.830	14.952	15.229	13.787
Personal cars, K_H	120	120	120	120
Total number of vehicles, K_V	130.830	134.952	135.229	133.787
Commercial transportation , $T_H + T_M$	7.467	10.309	10.500	10.500
Capacity, K_F	25	25	25	27.603
Capacity utilization, u	0.811	0.815	0.816	0.739

Looking at the table, the following conclusions can be made:

1. The first two scenarios do not even reach the required level of the commercial transportation services of 10.5 units in this case, because they do not incorporate the existing congestion and the network's effects. In the long run, it may lead to a slow-down in economic growth and development in the region.

2. While the third scenario does incorporate congestion and the network positive externality, it is not a social optimum since the scenario regards transportation fixed facilities as given or provided outside the private decision-making process. Under this scenario for given immobile capital K_F, the value of the capacity utilization u of 0.816 is sub-optimal.

3. The fourth scenario results in both, optimal network in terms of capacity and optimal capacity

utilization, given the existing economic conditions.

As already mentioned, the results are obtained in a static framework. Therefore, let us illustrate and explain the results based on economic theory. The following diagram compares scenario 1 with scenario 2:

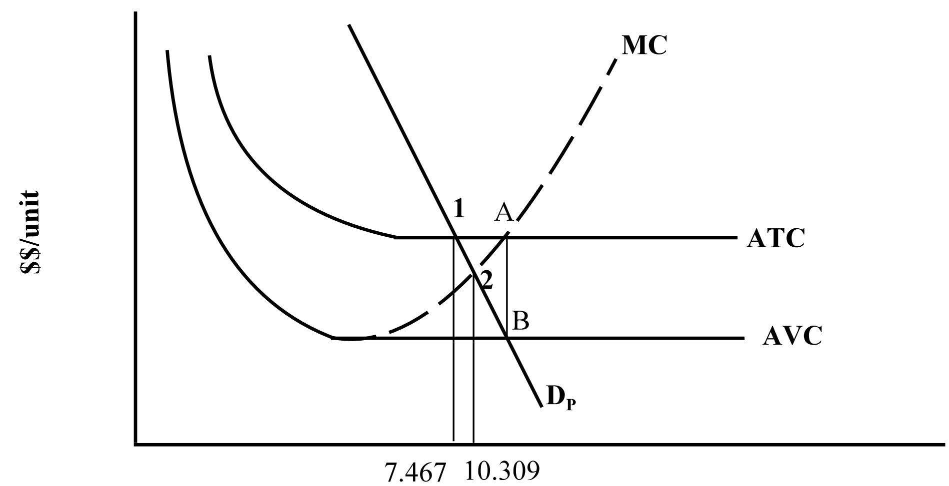

Fig. 3

In the above diagram, $\mathbf{D_P}$ is the private demand curve that does not take into account positive network externality. If private provider does not take into account congestion, then the relevant supply characteristic is average variable cost **AVC**. However, with congestion the relevant supply characteristic should be marginal cost **MC**. This is a standard representation of the congestion effect in transportation literature (see, for example, Button, 1993).

So, if congestion is not taken into account, the provider should choose point B. However at this point, there is congestion cost represented by the distance AB. If the provider believes that the **AVC** curve is of the shape shown in the diagram, then an imaginary congestion tax should be imposed to reflect the reality, which shifts the **AVC** curve upwards. The new line is marked **ATC** or average total cost because it includes effect of transportation immobile capital K_F through congestion tax, and in such a sense it is average total cost curve. The real output of 7.467 units in the simulation exercise is a result of intersection between the **D$_P$** curve and the **ATC** curve at point **1**.

On the other hand, if the provider recognized the existing congestion, it would choose point **2** with resulting output of 10.309 units. Hence, the difference between points **1** and **2** in figure 3 is actually the difference between scenario 1 and scenario 2. Now let us discuss the following diagram:

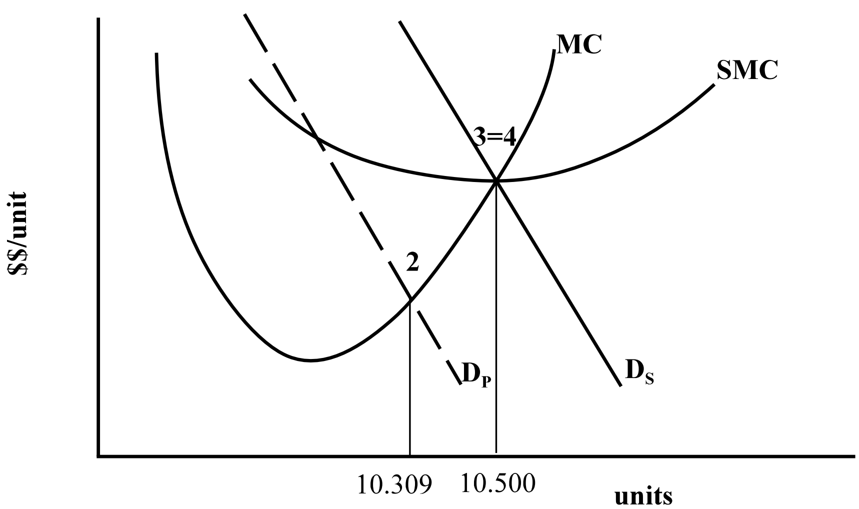

Fig. 4

In the above diagram, **D$_S$** is the social demand curve that takes into account positive indirect network externality, and **SMC** is the social marginal cost curve that incorporates cost of transportation facilities (immobile capital) as variable cost. Point **2** is the same as in Fig. 1. It is the result of intersection between the private demand **D$_P$** curve and the private marginal cost **MC** curve. Positive consumption externality shifts the private demand curve to its social demand curve **D$_S$** position. Point **3** is the result of intersection between the **MC** curve and the **D$_S$** curve, while point **4** is the result of intersection between the $\mathbf{D_S}$ and the **SMC** curve. In terms of production of transportation services, the two points coincide. However, social optimum at point **4** is efficient while private sub-optimum at point **3** is not since private provider does not produce at minimum total cost. Hence, points **2**,**3** and **4** correspond to scenarios 2, 3 and 4.

Conclusion

It appears to be that conventional interpretation of consumption and production of transportation services in transportation economic literature results in inefficient solution. In terms of the discussed framework, the conventional approach corresponds to scenario 1 or at best scenario 2. As we have already seen, both lead to inefficiencies at microeconomic level with undesired macroeconomic consequences.

In order to avoid these inefficiencies, transportation should be analysed as production and consumption of multiple transportation services, both freight and passenger, in a transportation network, the latter being a part of an economic system. Such a framework would help us to directly incorporate effects of transportation fixed facilities such as highways, terminals, airports, seaport and others as well as positive network externalities.

It is also worth mentioning, that positive network (consumption) externalities are of two types, direct and indirect. Direct externalities are associated with the *physical* network expansion which eventually leads to a better accessibility to new locations by more diverse groups of population. In turn, indirect network externalities are associated with an increase in economic activity around the network due to various reasons, for example, booming regional development, which increases the network's value for current users.

Even though in this paper only positive indirect network externalities are modelled and the framework is static, nevertheless it does provide a rationale for the future work in the field. The presented methodology appears more attractive since it closer resembles the reality of transportation in comparison with conventional approaches. As well, the described framework might become a building block for studying such transportation phenomena as sustainable transportation, intermodal transportation, intelligent transportation systems and others, which is almost impossible with the tools developed under conventional economics of transportation.

In the meantime, however, the following steps should be realized:

1. Demand for commercial transportation and personal transportation should be presented as the following functions:

$$T_H = f(Y, N) \text{ and } T_P = f(K_H , Y, N)$$

where N is the number of the network users, to reflect *direct* network externalities.

2. Depending on relative prices of personal T_P and commercial T_H transportation of households and their disposable income, a choice should be made between T_P and T_H (allocation decision).

3. A link between Block 0 and Block 2 in Fig. 1 should be developed since transportation capital K_C and K_F as well as transportation labour L_C and manufacturing capital K and labour L are related.

4. The affected sectors of the economy should include:

- construction (K_F , K);
- automobile (K_C, K_H);
- labour (L_C,,L).

References

Boyer, K.D. *Principles of Transportation Economics.* Addison Wesley Longman, Inc., 1998

Button, K.J. *Transport Economics. Second Edition.* Cambridge, UK: Edward Elgar, 1993

Cain, L.P. Historical Perspective on Infrastructure and US Economic Development. *Regional Science and Urban Economics,* No. 27, 1997, 117-138

Economides, N. The Economics of Networks. *International Journal of Industrial Organization,* Vol. 14, No. 2, 1996, 673 ff

Ferrara, E. and Marcellino, M. TFP, Costs and Public Infrastructure. Italy: Bocconi University and IGIER, 2000

Katz, M.L. and Shapiro, C. Systems Competition and Network Effects. *Journal of Economic Perspectives,* Vol. 8, 1994, 93-115

Krugman, P. *Development, Geography and Economic Theory.* Cambridge: MIT Press, 1995

Lipsey, R.G. and Steiner, P.O. *Economics*. New York: Harper and Row, 1969

Page, W.H. and Lopatka, J.E. Network Externalities. *University of Gent Encyclopaedia of Law and Economics,* 1997

Rohr, R. and Williams, I. Modelling the Regional Impacts of the Channel Tunnel. *Environment and Planning*, Vol. 21, 1994, 555-567

Roson, R. A General Equilibrium Analysis of the Italian Transport System. In Bannister, D., Capello, R. and Nijkamp, P. *European Transport and Communication Networks: Policy Evolution and Change.* New York: John Wiley and Sons, 1995

Sharkey, W.W. Network Models in Economics, 1993, mimeo

Tatom, J.A. Is an Infrastructure Crisis Lowering the Nation's Productivity? *Federal Reserve Bank of Saint Louis Review,* Vol. 75, No. 6, 1993, 1-19

Yevdokimov, Y. Economic Modeling of Congestion in a Transportation Network, *Proceedings of the 36th Annual Conference of Canadian Transportation Research Forum,* Vancouver, May 6-8, 2001, 705-720